MYNTRA

<hacker-ramp/>

Team Name: Sundar log

Team Details:

Shriyam Singh Tiwari
Archisha Srivastava
Jaspreet Kaur

PROBLEM STATEMENT

The fashion industry is highly dynamic and driven by ever-changing consumer preferences and social influences. Traditional trend forecasting is often slow and imprecise, making it hard for brands to keep up.

SOLUTION

Objective: Predict user preferences based on demographic data and current fashion trends using machine learning.

Steps Involved

1. Web Scraping:
    - Purpose: Collect relevant fashion data from online sources like blogs, social media, and e-commerce sites.
    - Tools Used: BeautifulSoup & Pandas.
    - Data Points Collected: Trends, popular products, user reviews, fashion articles.

2. Data Cleaning:
    - Purpose: Ensure collected **data** is accurate and usable.
    - Process:
        - Remove duplicates and irrelevant data.
        - Handle missing values.
        - Normalize data formats.
3. Preprocessing:
    - Purpose: Transform data into a suitable format for **machine learning**.
    - Process:
        - Feature Engineering: Create new features to enhance predictive power.
        - Scaling and Normalization: Ensure consistent data range.
        - Data Splitting: Divide data into training, validation, and test sets.

Model Training:

- Purpose: Train a machine learning model to predict user preferences.
- Process:
    - Model Selection: **Random Forest Classifier**
    - Training: Train the model on the training data.
    - Hyperparameter Tuning: Optimize parameters for better performance.
    - Evaluation: Assess performance using metrics like accuracy and F1-score.

Results

- Predictions: The model predicts user preferences based on **demographic** and **trend** data.
- Insights: Provides actionable insights for brands to understand e**merging trends** and tailor their offerings.

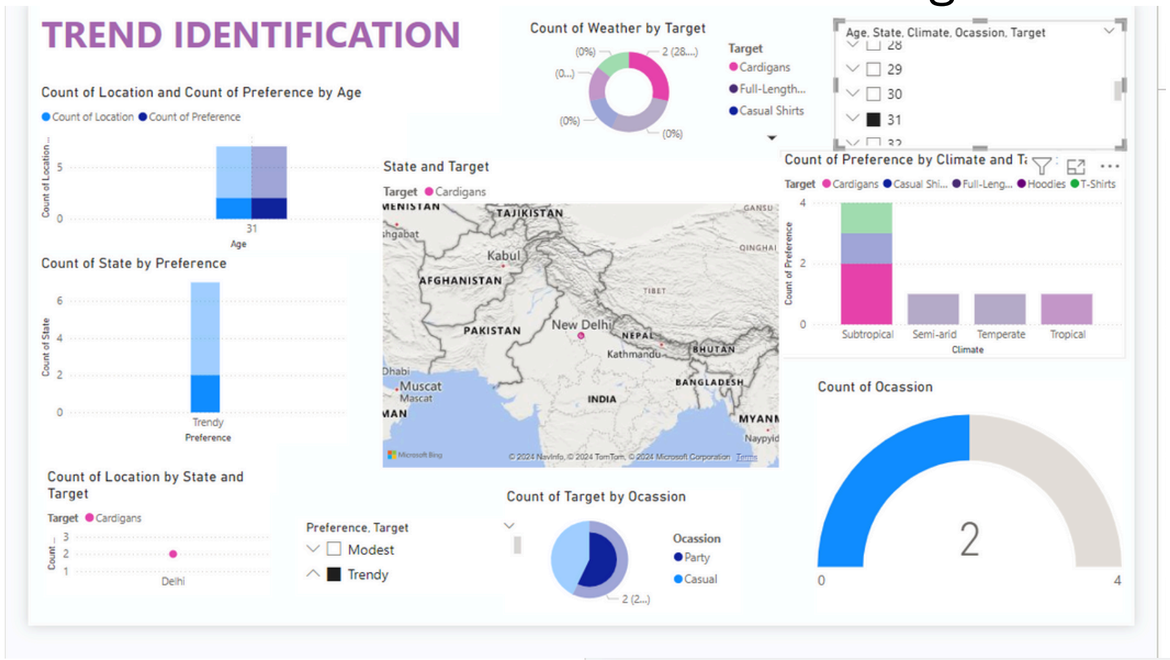
Data collection, Cleaning and Visualization

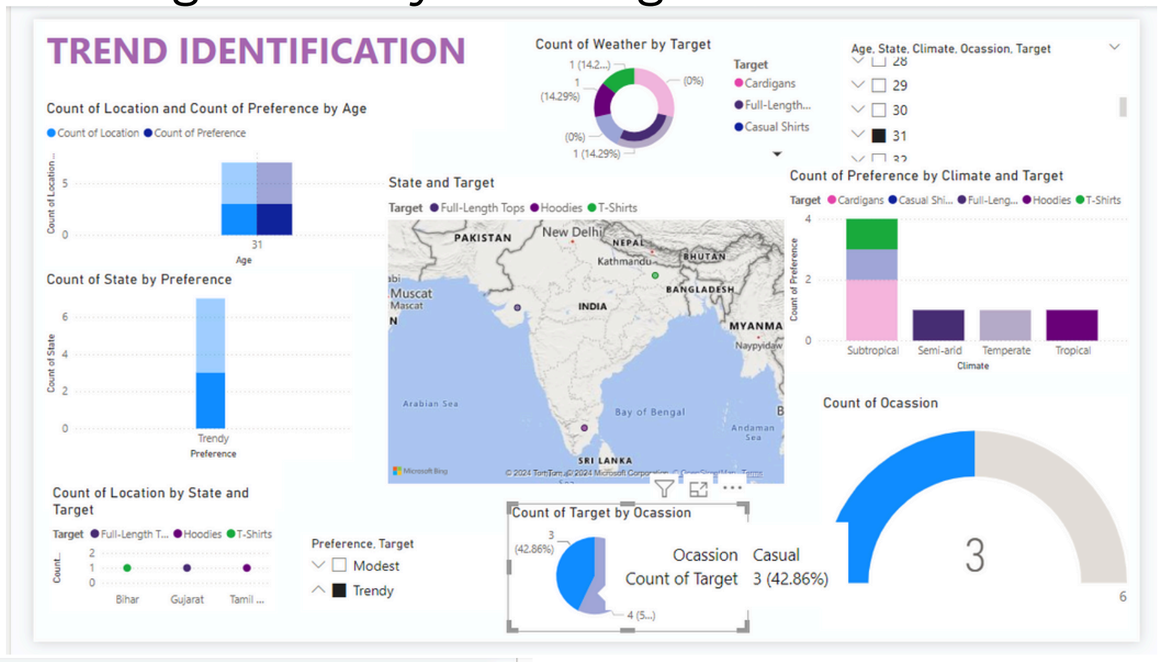
Data collection done from Amazon and Myntra ,done using BeautifulSoup & Pandas

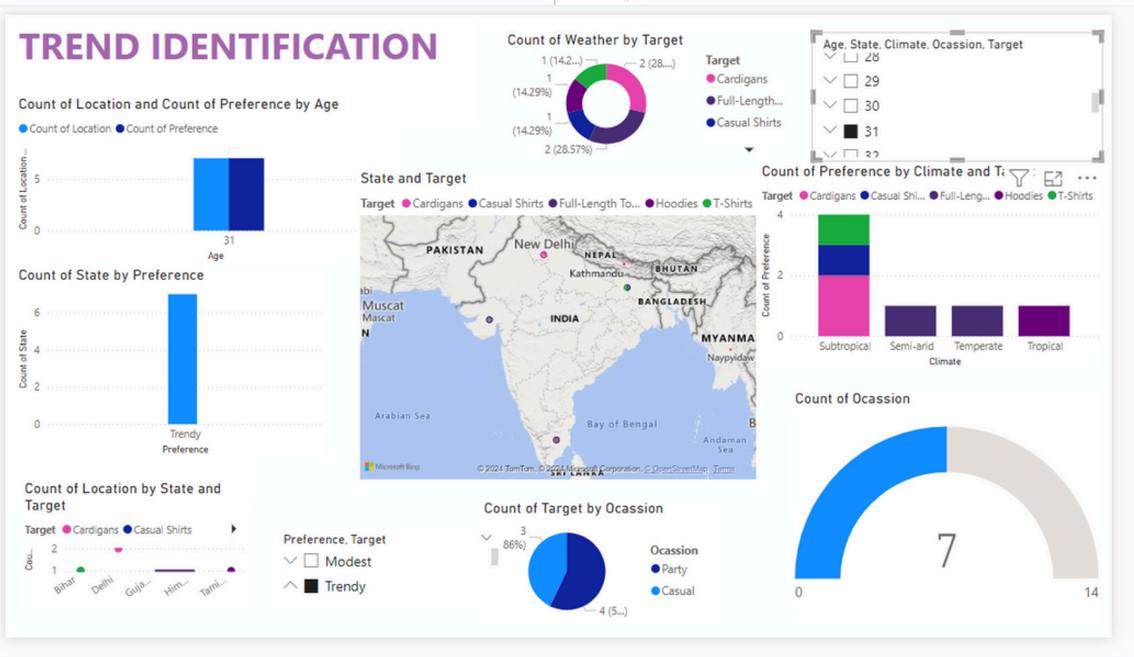
Cleaned Data stored in Train.csv

	Age	Climate	Weather	Location	Preference	State	Ocassion	Target
2	27	Semi-arid	Humid	Rural	Modest	Gujarat	Party	Short Kurtis
3	23	Tropical	Rainy	Suburban	Trendy	Tamil Nad	Casual	Short Kurtis
4	29	Subtropica	Cloudy	Urban	Modest	Bihar	Festival	Short Kurtis
5	26	Temperate	Sunny	Rural	Trendy	Himanchal	Casual	Short Kurtis
6	28	Semi-arid	Sunny	Urban	Modest	Delhi	Party	Short Kurtis
7	24	Semi-arid	Humid	Rural	Trendy	Gujarat	Casual	Short Kurtis
8	30	Tropical	Rainy	Suburban	Modest	Tamil Nad	Festival	Short Kurtis
9	32	Subtropica	Cloudy	Urban	Trendy	Bihar	Casual	Short Kurtis
10	27	Temperate	Sunny	Rural	Modest	Himanchal	Party	Short Kurtis
11	25	Semi-arid	Sunny	Urban	Trendy	Delhi	Casual	Short Kurtis
12	23	Semi-arid	Humid	Rural	Modest	Gujarat	Festival	Short Kurtis
13	26	Tropical	Rainy	Suburban	Trendy	Tamil Nad	Casual	Short Kurtis
14	29	Subtropica	Cloudy	Urban	Modest	Bihar	Party	Short Kurtis
15	24	Temperate	Sunny	Rural	Trendy	Himanchal	Casual	Short Kurtis
16	28	Semi-arid	Sunny	Urban	Modest	Delhi	Party	Short Kurtis
17	30	Subtropica	Sunny	Urban	Modest	Delhi	Casual	Cardigans
18	32	Subtropica	Humid	Rural	Trendy	Gujarat	Party	Cardigans
19	28	Tropical	Rainy	Suburban	Modest	Tamil Nad	Casual	Cardigans
20	35	Subtropica	Cloudy	Urban	Trendy	Bihar	Festival	Cardigans
21	33	Temperate	Sunny	Rural	Modest	Himanchal	Casual	Cardigans
22	31	Subtropica	Sunny	Urban	Trendy	Delhi	Party	Cardigans
23	29	Tropical	Humid	Rural	Modest	Gujarat	Casual	Cardigans
24	34	Subtropica	Rainy	Suburban	Trendy	Tamil Nad	Party	Cardigans
25	36	Temperate	Cloudy	Urban	Modest	Bihar	Festival	Cardigans
26	32	Subtropica	Sunny	Rural	Trendy	Himanchal	Casual	Cardigans
27	30	Subtropica	Sunny	Urban	Modest	Delhi	Party	Cardigans
28	27	Tropical	Humid	Rural	Trendy	Gujarat	Casual	Cardigans
29	31	Subtropica	Rainy	Suburban	Modest	Tamil Nad	Festival	Cardigans

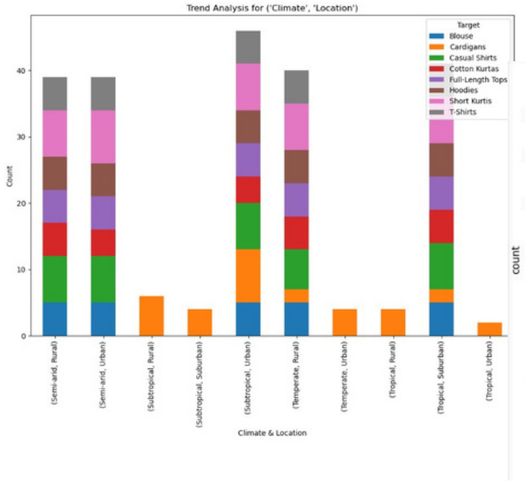
Dashboard creation using MS PowerBi. This gives analytical insights to the data

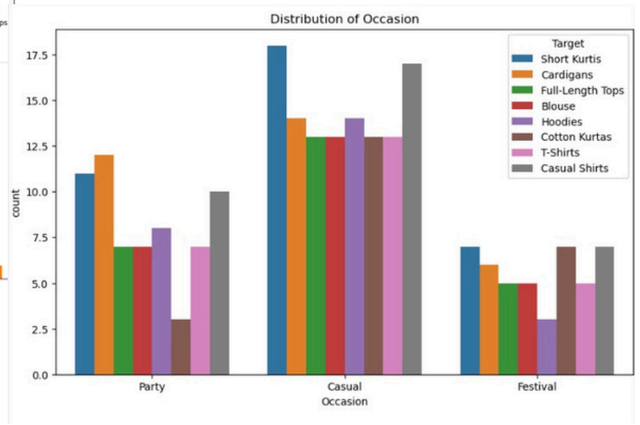


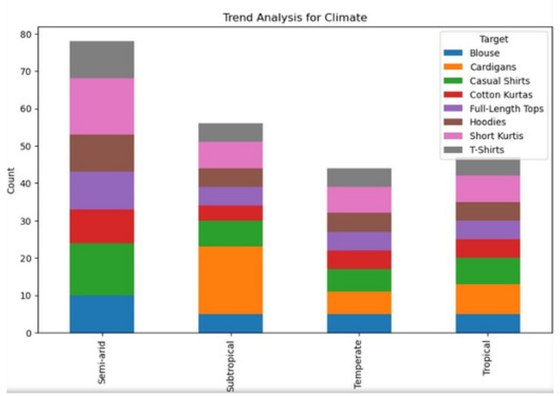


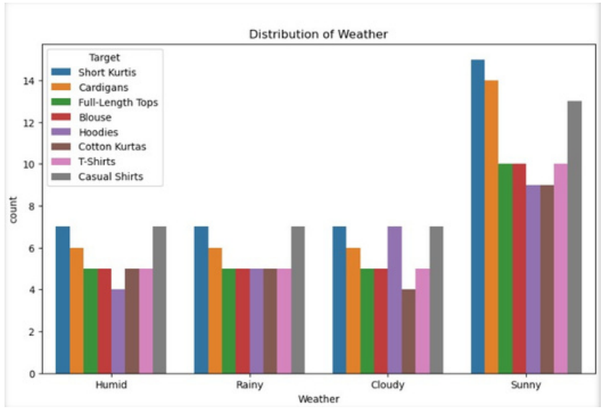


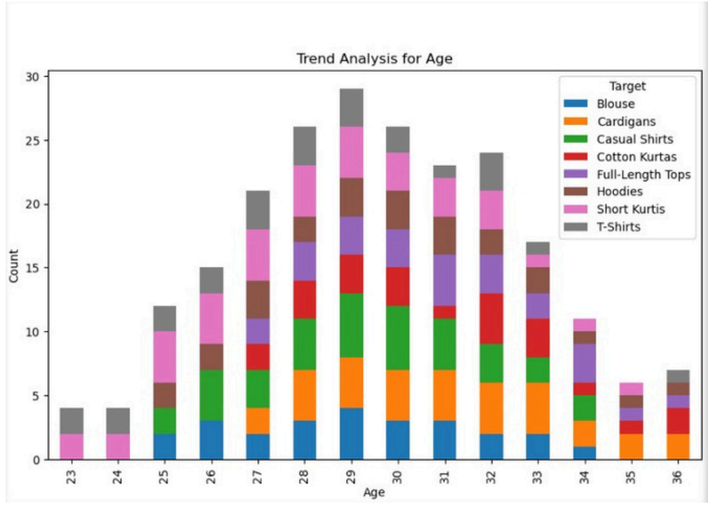
Model training

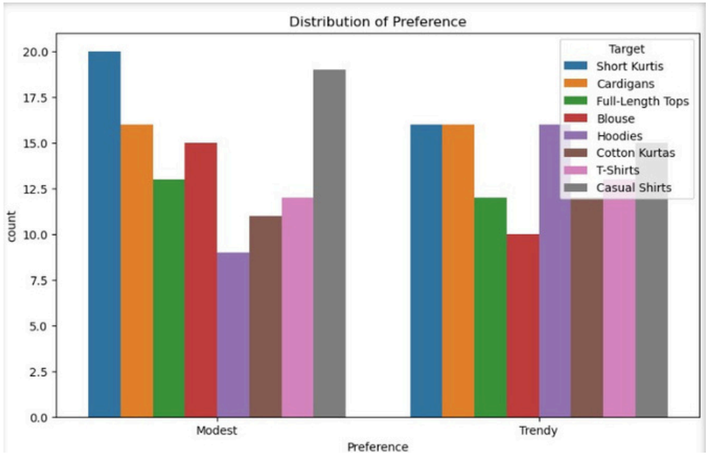


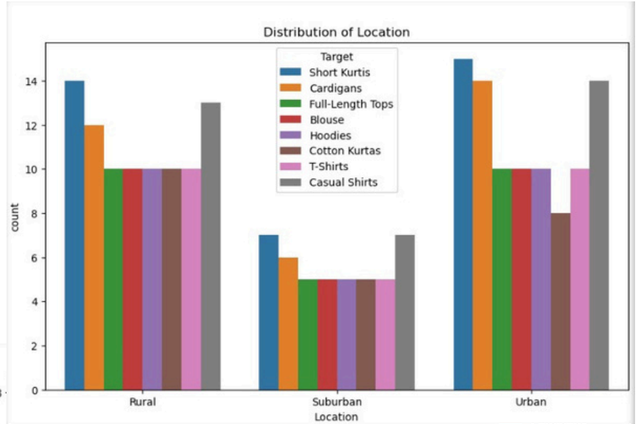


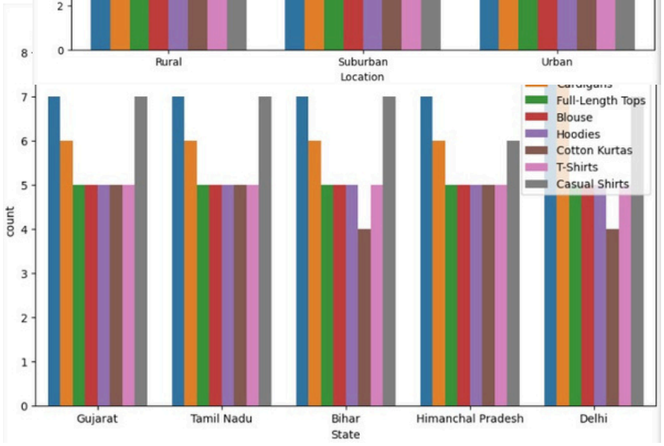




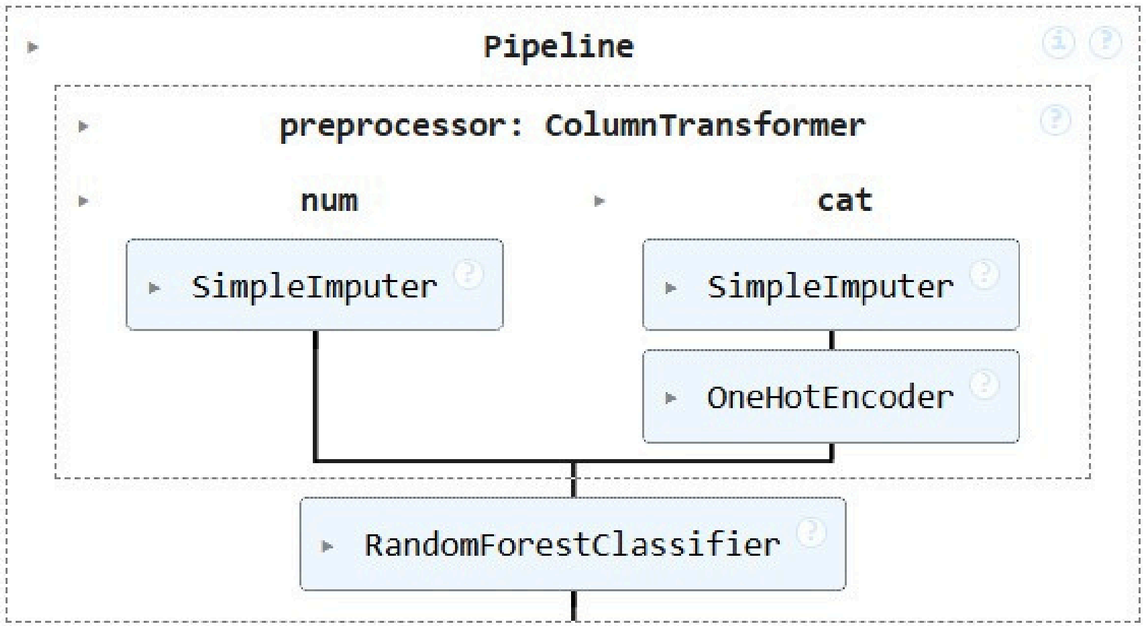








[70]:


Pipeline
preprocessor: ColumnTransformer
num
cat
SimpleImputer
SimpleImputer
OneHotEncoder
RandomForestClassifier

Model Output:

```
First few rows of the prediction results:
          Actual     Predicted
0  Short Kurtis  Short Kurtis
1  Short Kurtis  Short Kurtis
2  Short Kurtis      T-Shirts
3  Short Kurtis  Short Kurtis
4  Short Kurtis  Short Kurtis
```